RSAC | 2025 Conference

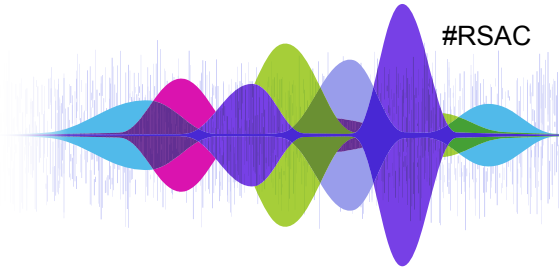
Many Voices.
One Community.

SESSION ID: USENIX-M01B

Smudged Fingerprints: Characterizing and Improving the Performance of Web Application Fingerprinting

Brian Kondracki

Cybersecurity Analyst
Jane Street Capital
https://www.linkedin.com/in/briankondracki/



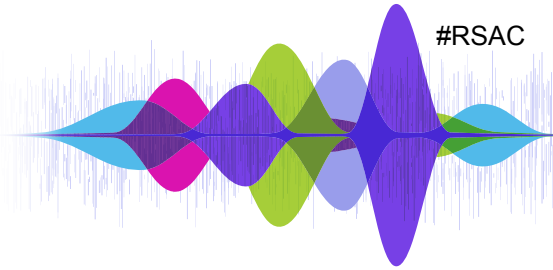
Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the presenters individually and, unless expressly stated to the contrary, are not the opinion or position of RSA Conference LLC or any other co-sponsors. RSA Conference LLC does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented.

Attendees should note that sessions may be audio- or video-recorded and may be published in various media, including print, audio and video formats without further notice. The presentation template and any media capture are subject to copyright protection.

© 2025 RSA Conference LLC or its affiliates. The RSAC and RSAC CONFERENCE logos and other trademarks are proprietary. All rights reserved.

Web Applications on the Internet



Open source web applications are the heart of the Internet

Web Applications on the Internet

Open source web applications are the heart of the Internet

Web Applications on the Internet

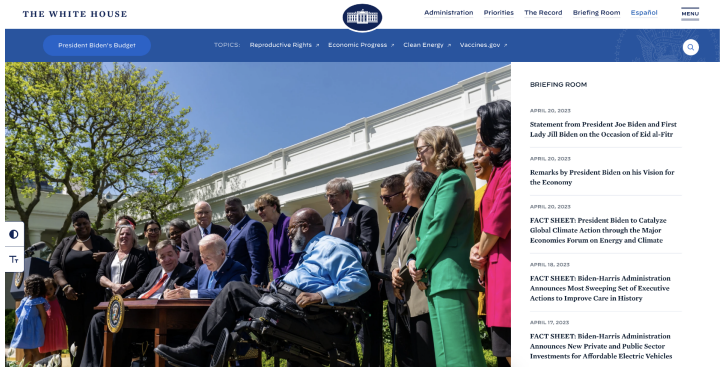
Open source web applications are the heart of the Internet

Web Applications on the Internet

Open source web applications are the heart of the Internet

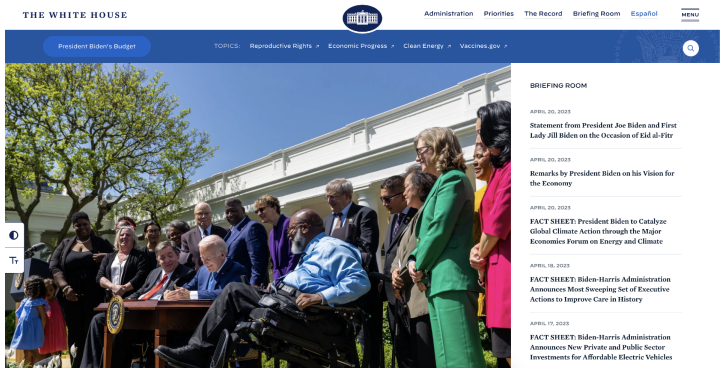
Web Applications on the Internet

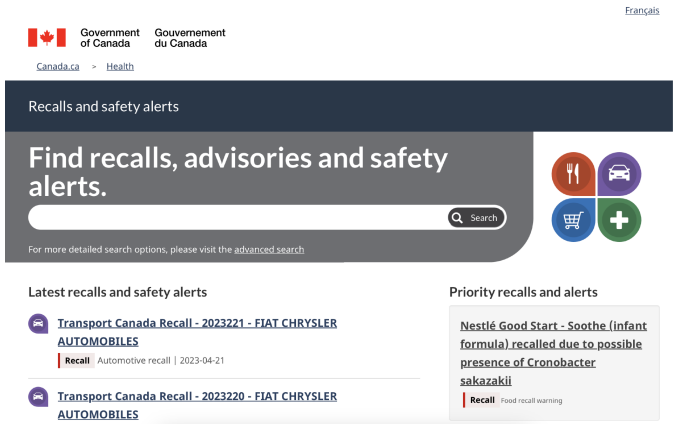
Open source web applications are the heart of the Internet


THE WHITE HOUSE
Administration
Priorities
The Record
Briefing Room
Español
MENU
President Biden's Budget
TOPICS:
Reproductive Rights
Economic Progress
Clean Energy
Vaccines.gov
BRIEFING ROOM
APRIL 20, 2023
Statement from President Joe Biden and First Lady Jill Biden on the Occasion of Eid al-Fitr
APRIL 20, 2023
Remarks by President Biden on his Vision for the Economy
APRIL 20, 2023
FACT SHEET: President Biden to Catalyze Global Climate Action through the Major Economies Forum on Energy and Climate
APRIL 18, 2023
FACT SHEET: Biden-Harris Administration Announces Most Sweeping Set of Executive Actions to Improve Care in History
APRIL 17, 2023
FACT SHEET: Biden-Harris Administration Announces New Private and Public Sector Investments for Affordable Electric Vehicles

Web Applications on the Internet

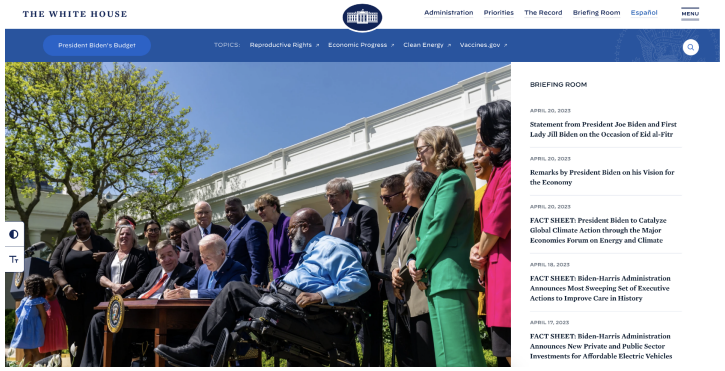
Open source web applications are the heart of the Internet

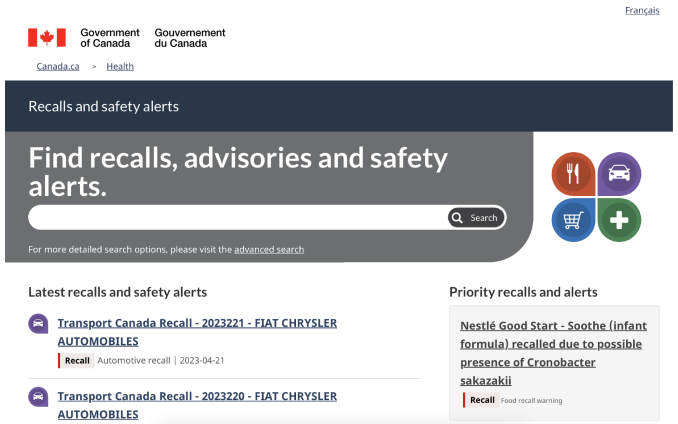

THE WHITE HOUSE
Administration
Priorities
The Record
Briefing Room
Español
MENU
President Biden's Budget
TOPICS:
Reproductive Rights
Economic Progress
Clean Energy
Vaccines.gov
BRIEFING ROOM
APRIL 20, 2023
Statement from President Joe Biden and First Lady Jill Biden on the Occasion of Eid al-Fitr
APRIL 20, 2023
Remarks by President Biden on his Vision for the Economy
APRIL 20, 2023
FACT SHEET: President Biden to Catalyze Global Climate Action through the Major Economies Forum on Energy and Climate
APRIL 18, 2023
FACT SHEET: Biden-Harris Administration Announces Most Sweeping Set of Executive Actions to Improve Care in History
APRIL 17, 2023
FACT SHEET: Biden-Harris Administration Announces New Private and Public Sector Investments for Affordable Electric Vehicles

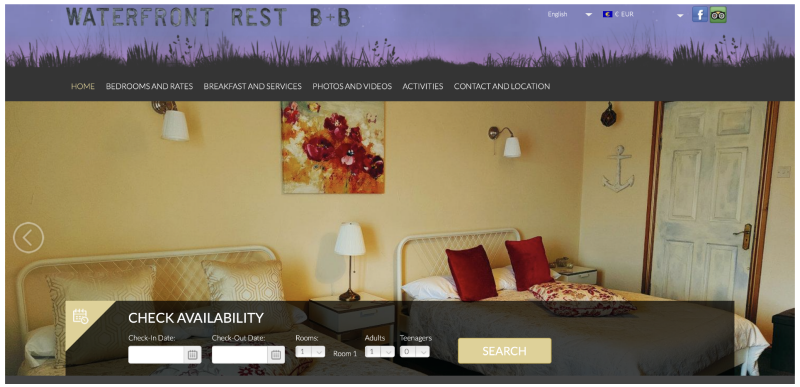

Français
Government of Canada
Gouvernement du Canada
Canada.ca > Health
Recalls and safety alerts
Find recalls, advisories and safety alerts.
Search
For more detailed search options, please visit the advanced search
Latest recalls and safety alerts
Transport Canada Recall - 2023221 - FIAT CHRYSLER AUTOMOBILES
Recall Automotive recall | 2023-04-21
Transport Canada Recall - 2023220 - FIAT CHRYSLER AUTOMOBILES
Priority recalls and alerts
Nestlé Good Start - Soothe (infant formula) recalled due to possible presence of Cronobacter sakazakii
Recall Food recall warning

Web Applications on the Internet

Open source web applications are the heart of the Internet


THE WHITE HOUSE
Administration
Priorities
The Record
Briefing Room
Español
MENU
President Biden's Budget
TOPICS:
Reproductive Rights
Economic Progress
Clean Energy
Vaccines.gov
BRIEFING ROOM
APRIL 20, 2023
Statement from President Joe Biden and First Lady Jill Biden on the Occasion of Eid al-Fitr
APRIL 20, 2023
Remarks by President Biden on his Vision for the Economy
APRIL 20, 2023
FACT SHEET: President Biden to Catalyze Global Climate Action through the Major Economies Forum on Energy and Climate
APRIL 18, 2023
FACT SHEET: Biden-Harris Administration Announces Most Sweeping Set of Executive Actions to Improve Care in History
APRIL 17, 2023
FACT SHEET: Biden-Harris Administration Announces New Private and Public Sector Investments for Affordable Electric Vehicles


Français
Government of Canada
Gouvernement du Canada
Canada.ca > Health
Recalls and safety alerts
Find recalls, advisories and safety alerts.
Search
For more detailed search options, please visit the advanced search
Latest recalls and safety alerts
Transport Canada Recall - 2023221 - FIAT CHRYSLER AUTOMOBILES
Recall Automotive recall | 2023-04-21
Transport Canada Recall - 2023220 - FIAT CHRYSLER AUTOMOBILES
Priority recalls and alerts
Nestlé Good Start - Soothe (infant formula) recalled due to possible presence of Cronobacter sakazakii
Recall Food recall warning



The Consequences of Vulnerable Web Applications

Hackers are actively probing millions of WordPress sites

By Sergiu Gatlan — November 17, 2020 — 05:02 PM — 1

Unknown threat actors are scanning for WordPress websites with Epsilo
on over 150,000 sites and vulnerable to Function Injection attacks that c

IMAGE CREDITS: BRYCE DURBIN / TECHCRUNCH

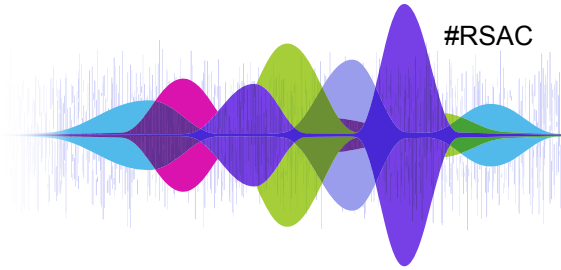
SECURITY

Hackers ar
WordPress sites to push
Windows and Mac malware

Lorenzo Franceschi-Bicchierai – 2:03 PM PST · January 29, 2025

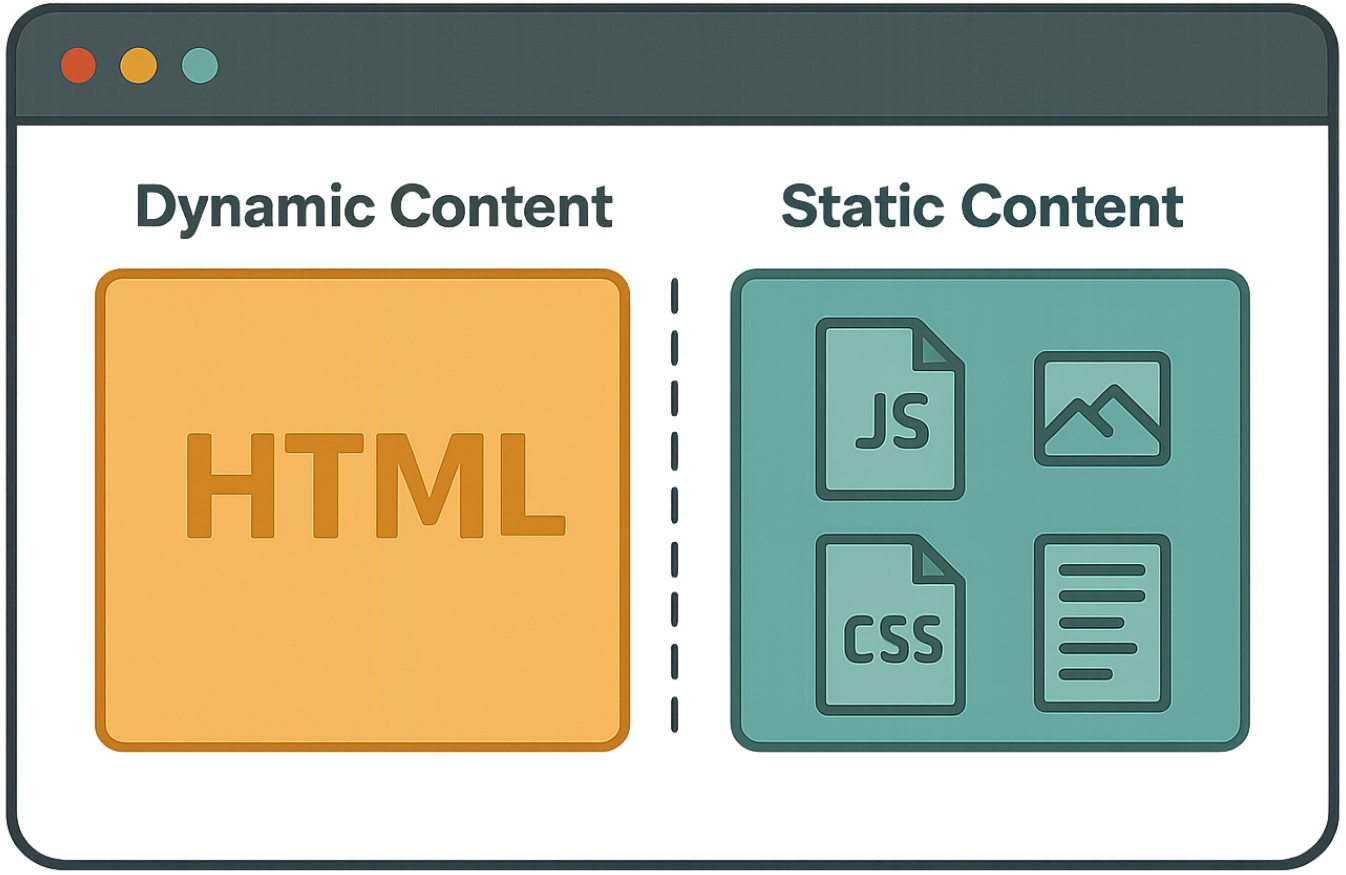
Drupal Warns Web Admins to Update CMS Sites to Patch a Critical Flaw

Dec 19, 2019 — Swati Khandelwal



Effective web application fingerprinting is a powerful tool!

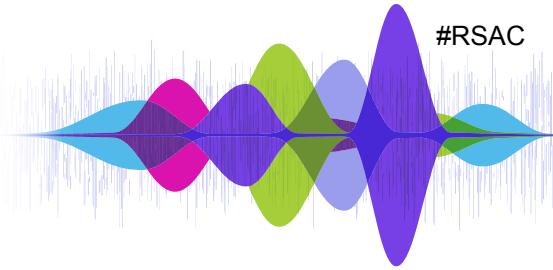
Anatomy of a Web Application



Dynamic Content Fingerprinting

```
<link href='https://example.com/wp-content/...' />
    ...
<meta name="generator" content="WordPress 5.8" />
    ...
<p>
    Powered by
    <a href="https://wordpress.org">WordPress</a>
</p>
```

Static Content Fingerprinting



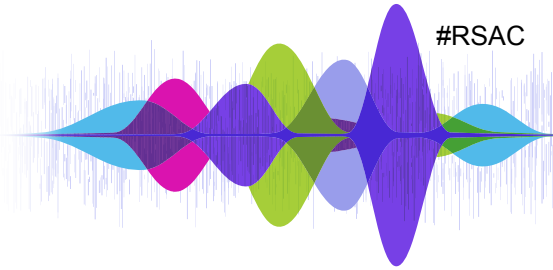
WordPress 5.8 - a380109124ae1089684a00c8ac9c5c68

```
   1   /******/ (function(modules) { // webpackBootstrap
           ...
1412               posts_per_page: 40
           ...
1708   /******/ });
```

WordPress 5.8.1 - 3086430ffe51587d0ac5ae51b4404df7

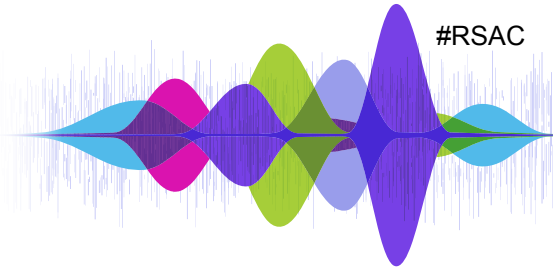
```
   1   /******/ (function(modules) { // webpackBootstrap
           ...
1412               posts_per_page: 80
           ...
1708   /******/ });
```

Research Questions

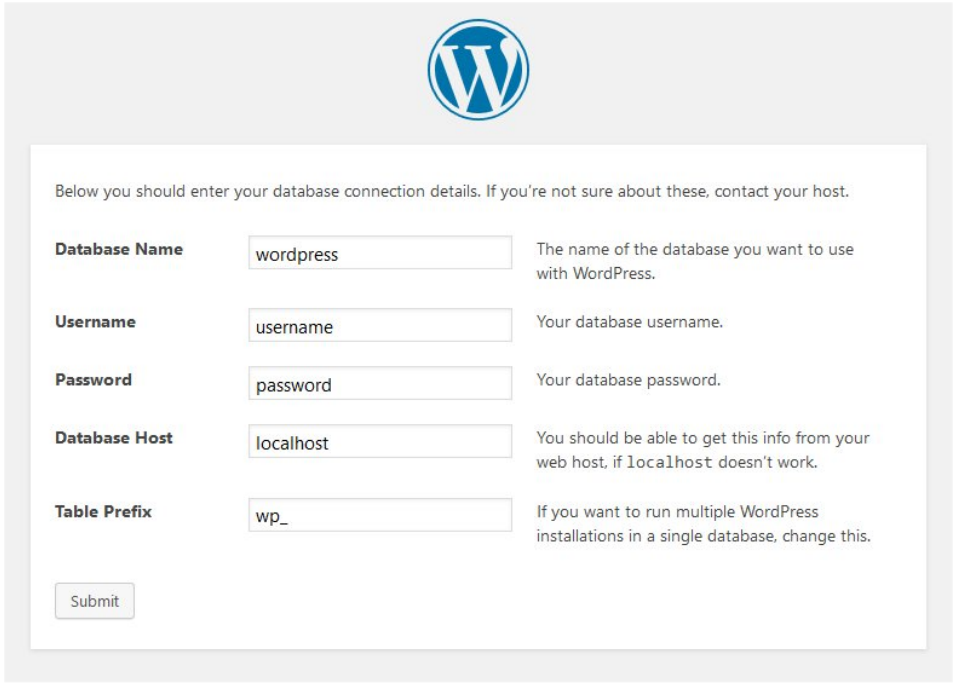


1. How effective are modern web application fingerprinting techniques?
2. How robust are these techniques to the noise added by real-world web application deployments?
3. What factors influence the performance of web application fingerprinting techniques in the wild? Can we address them?

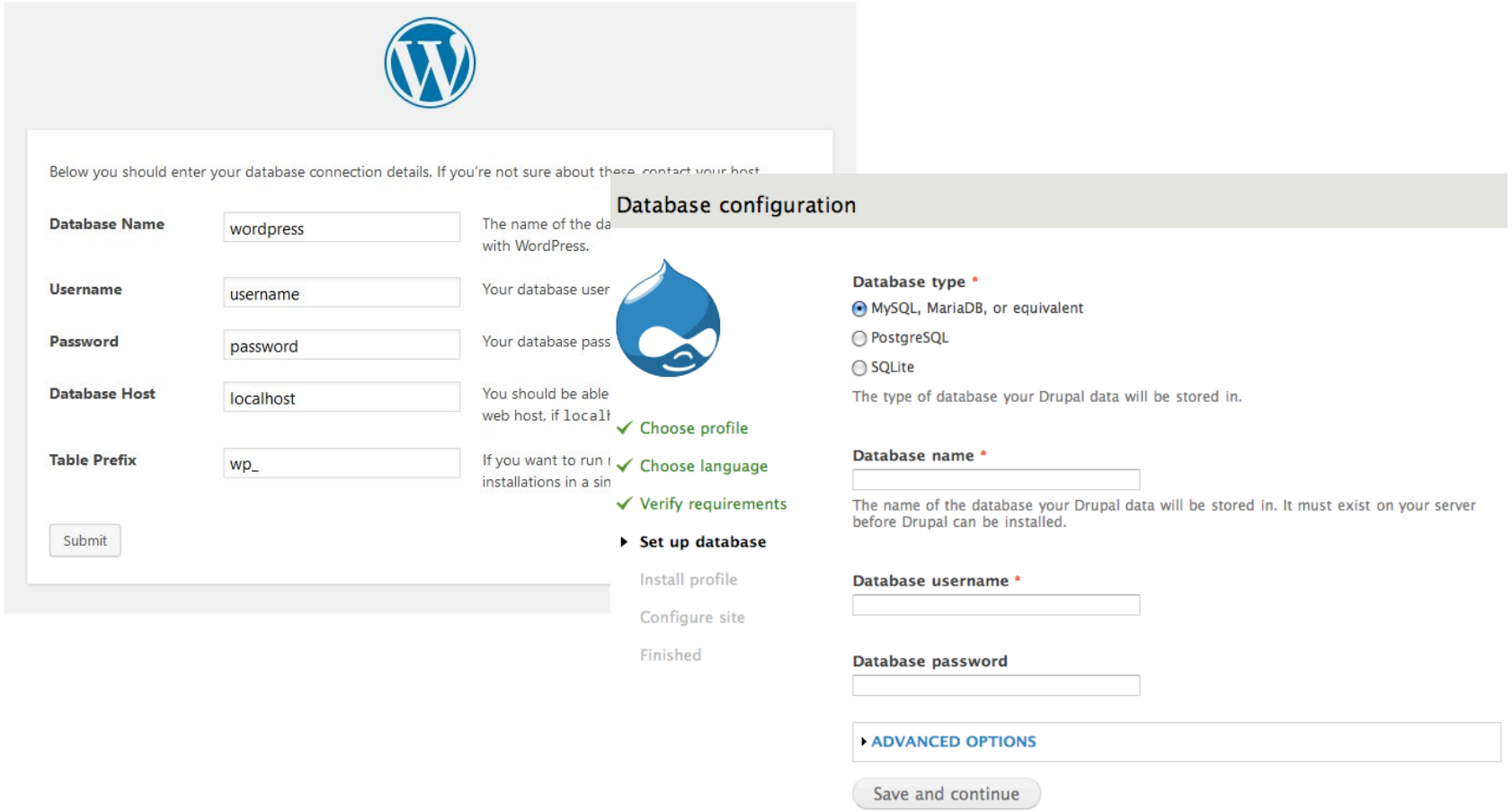
Web Application Installation Process



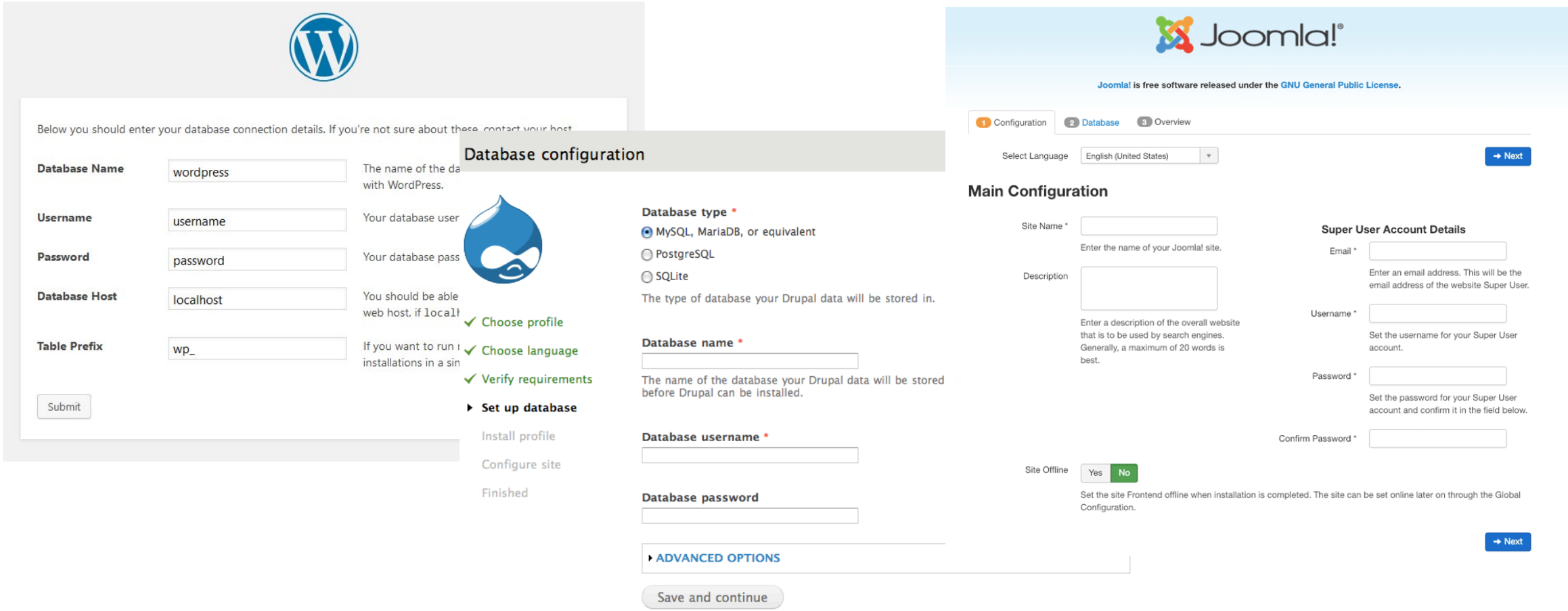
Web Application Installation Process


Below you should enter your database connection details. If you're not sure about these, contact your host.
Database Name
wordpress
The name of the database you want to use with WordPress.
Username
username
Your database username.
Password
password
Your database password.
Database Host
localhost
You should be able to get this info from your web host, if localhost doesn't work.
Table Prefix
wp_
If you want to run multiple WordPress installations in a single database, change this.
Submit

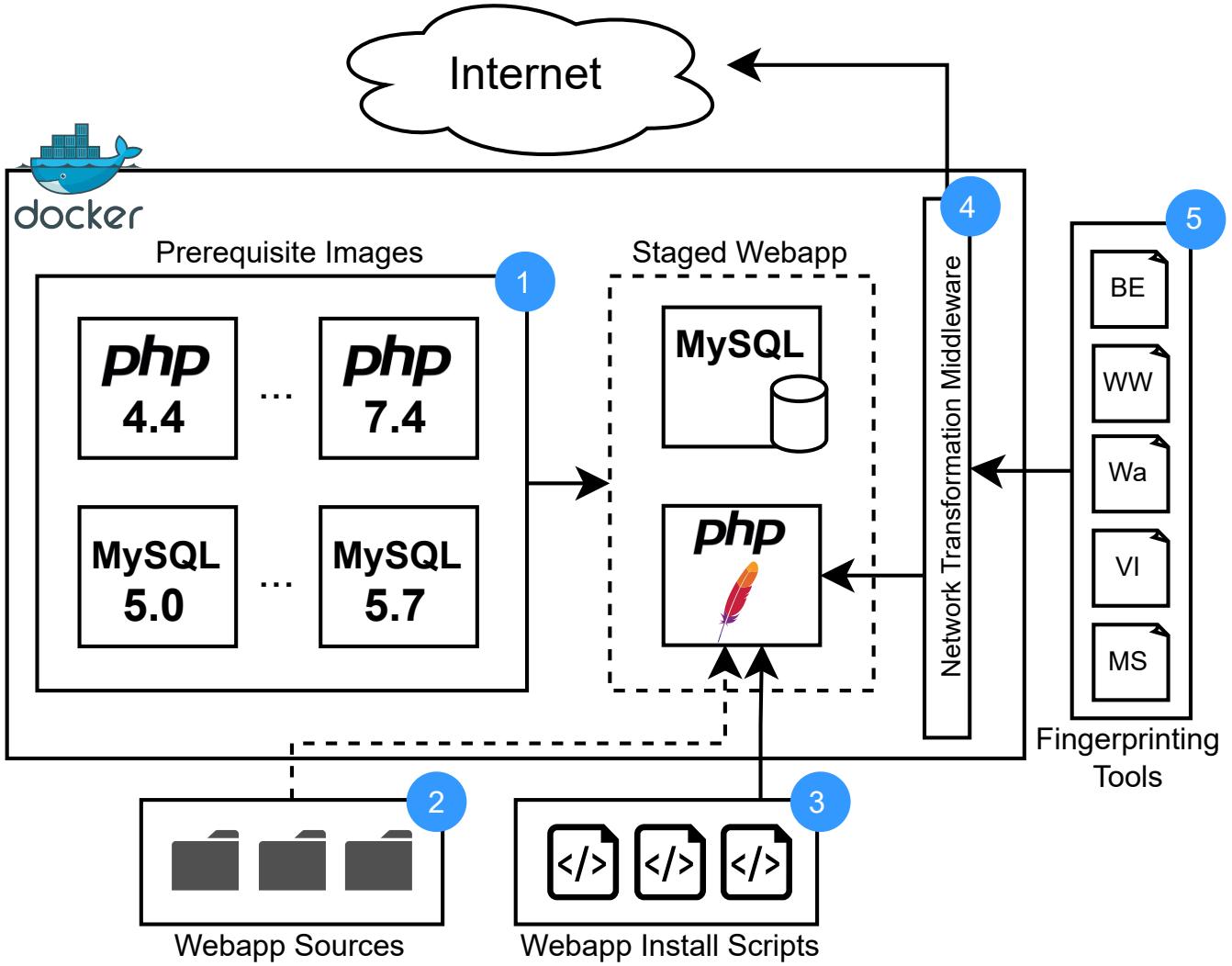
Web Application Installation Process



Web Application Installation Process


Below you should enter your database connection details. If you're not sure about these, contact your host.
Database Name
wordpress
The name of the da
with WordPress.
Username
username
Your database user
Password
password
Your database pass
Database Host
localhost
You should be able
web host, if local
Table Prefix
wp_
If you want to run
installations in a sin
Submit
Database configuration
Choose profile
Choose language
Verify requirements
Set up database
Install profile
Configure site
Finished
Database type *
MySQL, MariaDB, or equivalent
PostgreSQL
SQLite
The type of database your Drupal data will be stored in.
Database name *
The name of the database your Drupal data will be stored before Drupal can be installed.
Database username *
Database password
ADVANCED OPTIONS
Save and continue
Joomla!
Joomla! is free software released under the GNU General Public License.
1 Configuration
2 Database
3 Overview
Select Language
English (United States)
Next
Main Configuration
Site Name *
Enter the name of your Joomla! site.
Description
Enter a description of the overall website that is to be used by search engines. Generally, a maximum of 20 words is best.
Super User Account Details
Email *
Enter an email address. This will be the email address of the website Super User.
Username *
Set the username for your Super User account.
Password *
Set the password for your Super User account and confirm it in the field below.
Confirm Password *
Site Offline
Yes
No
Set the site Frontend offline when installation is completed. The site can be set online later on through the Global Configuration.
Next

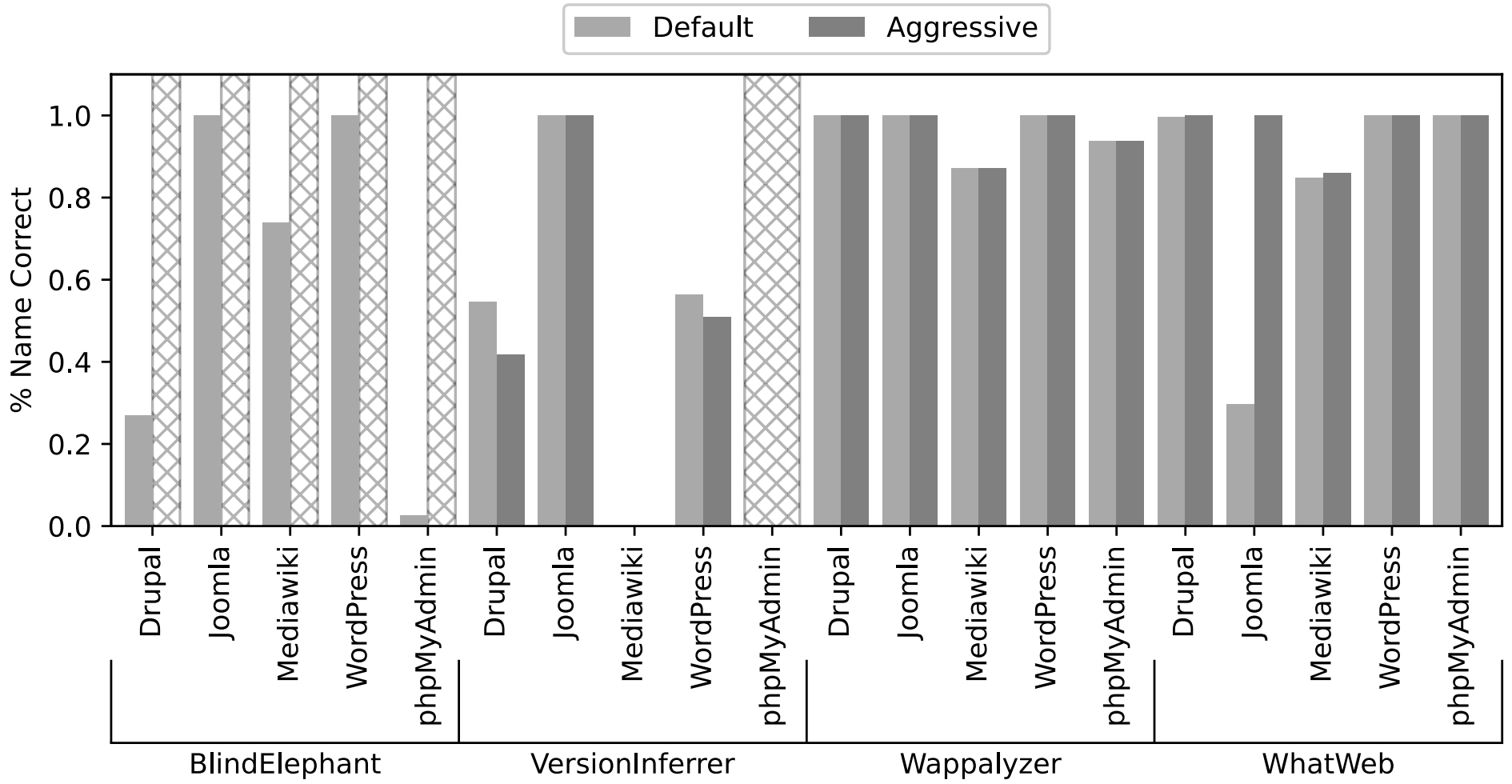
WASABO (Web Application SAndBOx)



Web Application Fingerprinting Tools

Tool	Fingerprinting Type		# Requests	
	Dynamic	Static	Default	Aggressive
BlindElephant	○	●	$10^0/10^1$	$10^0/10^4$
WhatWeb	●	●	10^0	10^2
VersionInferer	●	●	10^0	10^1
Wappalyzer	●	○	10^0	10^0
Metasploit-Joomla	●	○	10^0	-
Metasploit-WordPress	●	○	10^2	-

Web Application Type Prediction



Web Application Version Prediction

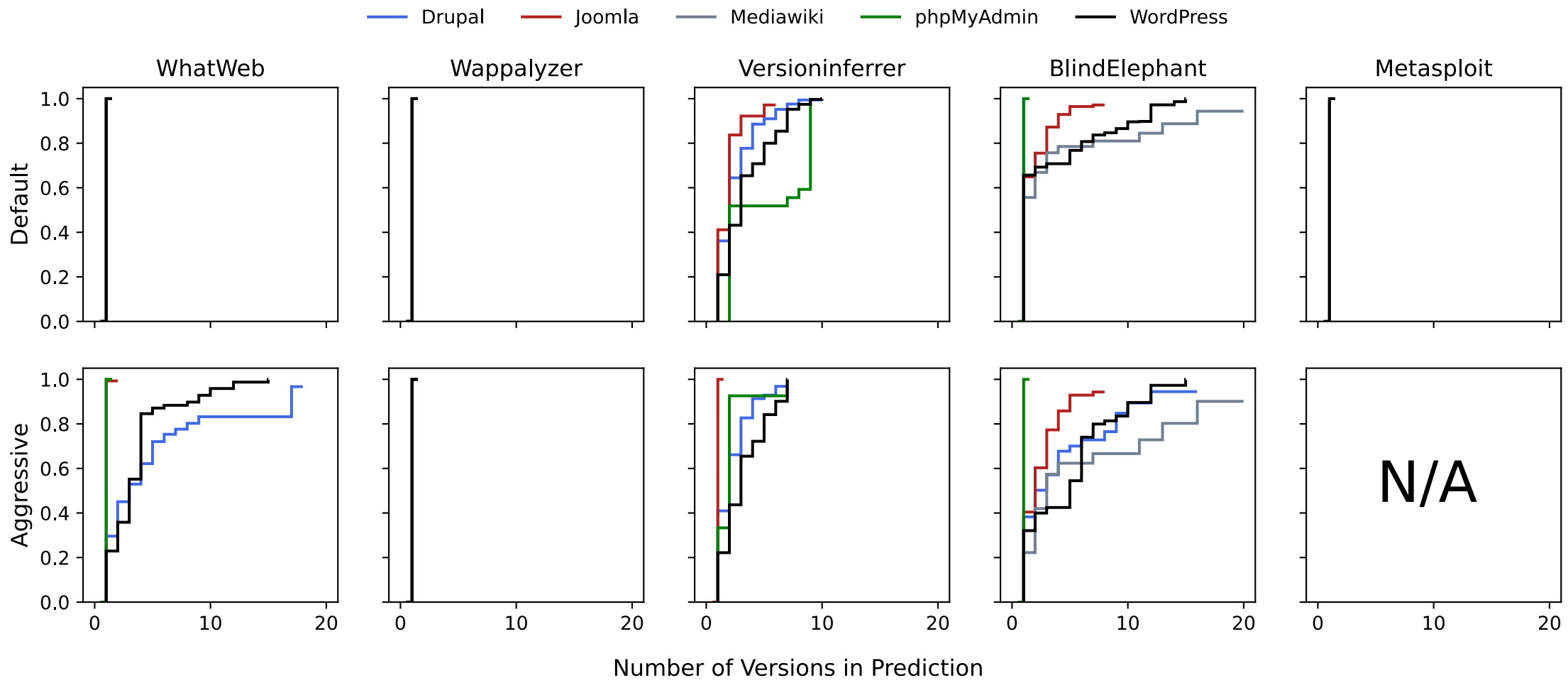
No Prior Knowledge

Tool	Webapp	Type	Version Matched (%) Major	Minor	Full
VersionInferrer	Drupal	D	53.6	53.6	53.6
		A	40.8	40.8	40.8
	Joomla	D,A	100.0	100.0	100.0
	Mediawiki	D,A	0.0	0.0	0.0
	WordPress	D	56.5	56.5	55.9
		A	50.9	50.9	50.2
	phpMyAdmin	D,A	0.0	0.0	0.0
Wappalyzer	Drupal	D,A	96.4	8.2	0.0
	Joomla	D,A	0.0	0.0	0.0
	Mediawiki	D,A	87.3	87.3	87.3
	WordPress	D,A	100.0	100.0	100.0
	phpMyAdmin	D,A	0.0	0.0	0.0
WhatWeb	Drupal	D	0.0	0.0	0.0
		A	97.4	72.4	65.5
	Joomla	D	0.0	0.0	0.0
		A	100.0	100.0	100.0
	Mediawiki	D	46.1	46.1	46.1
		A	46.7	46.7	46.7
	WordPress	D,A	100.0	100.0	100.0
	phpMyAdmin	D	31.2	31.2	31.2
		A	34.9	34.9	34.9

Prior Knowledge

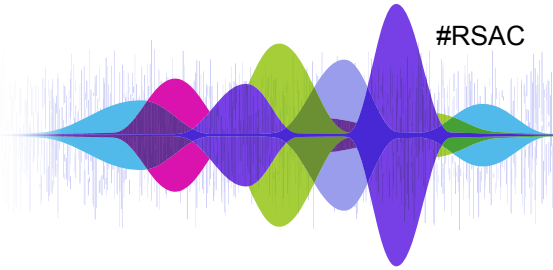
Tool	Webapp	Type	Version Matched (%) Major	Minor	Full
BlindElephant	Drupal	D	18.8	18.8	18.8
		A	71.4	71.4	70.4
	Joomla	D,A	100.0	100.0	100.0
	Mediawiki	D,A	98.2	97.6	95.2
	WordPress	D,A	100.0	100.0	93.9
	phpMyAdmin	D,A	54.2	54.2	54.2
Metasploit	Joomla	D	100.0	100.0	100.0
	WordPress	D	100.0	100.0	100.0

Version Mismatch - Collisions



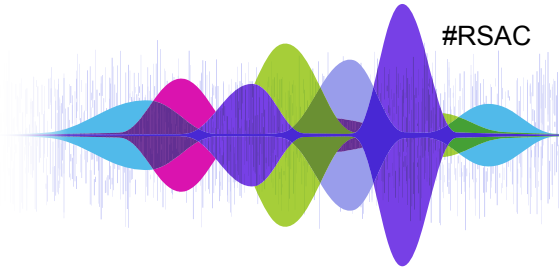
#RSAC
RSAC 2025 Conference

Real-world Fingerprinting Analysis



- Source ground truth from “showcase” pages on each web application’s website
  - 726 total websites discovered
- Determine baseline performance of each tool against all sites
- Repeat scans with series of transformations applied to network traffic of each tool

WASABO Network Middleware Transformations



Cache Breaker

Append a cache-breaking variable to all `GET` requests

Real Browser Requests

Intercept and replay all requests with a ChromeDriver instance

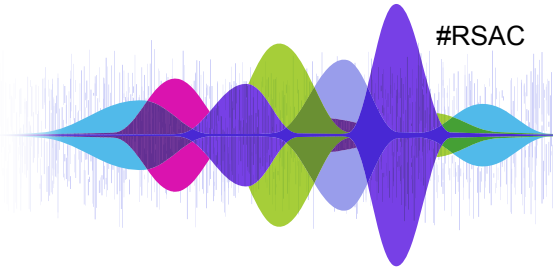
Path Prediction

Use paths listed on website homepage to identify common directory structure, modify subsequent request paths to match

Real-world Fingerprinting Performance

			Real-world Fingerprinting Accuracy (%)						
Tool	Webapp	Lab Accuracy (%)	Default	Cache Break	Path Predictor	ChromeDriver	Combined	Δ	Imp. Factor (%)
BlindElephant	Drupal	26.97	10.00	12.94	12.94	12.94	**13.02**	3.02	30.18
	Joomla	100.00	29.74	30.17	32.76	32.76	**33.77**	4.02	13.53
	Mediawiki	73.94	15.79	16.84	15.79	15.79	**17.89**	2.11	13.33
	WordPress	100.00	46.52	56.09	56.52	57.39	**59.74**	13.22	28.41
VersionInferrer	Drupal	54.60	30.00	31.76	30.00	36.47	**42.01**	12.01	40.04
	Joomla	100.00	12.07	12.50	12.93	18.10	**23.81**	11.74	97.28
	Mediawiki	00.00	3.16	**6.32**	3.16	4.21	**6.32**	3.16	100.00
	WordPress	56.45	39.39	40.00	40.87	53.04	**62.34**	22.94	58.24
Wappalyzer	Drupal	100.00	**73.96**	**73.96**	70.41	71.60	71.60	-2.37	-3.20
	Joomla	100.00	60.17	58.44	61.90	**66.23**	**66.23**	6.06	10.07
	Mediawiki	87.27	90.53	90.53	90.53	90.53	**91.58**	1.05	1.16
	WordPress	100.00	67.53	66.67	68.83	68.83	**70.13**	2.60	3.85
WhatWeb	Drupal	99.67	58.58	57.40	58.58	**59.76**	59.17	0.59	1.01
	Joomla	29.78	18.18	19.91	18.18	19.48	**21.21**	3.03	16.67
	Mediawiki	84.84	85.26	85.26	85.26	89.47	**90.53**	5.26	6.17
	WordPress	100.00	65.37	64.94	65.37	69.26	**69.70**	4.33	6.62

Key Takeaways - Apply what you learned



- Fingerprinting tools perform well in “best-case” scenario
- Tools often rely on web applications remaining unchanged from training
- WASABO can be utilized to improve performance of *any* tool with minimal setup
- New tools and techniques are still required to reliably identify deployments online
  - Secure online infrastructure
  - Study Internet-wide trends

RSAC 2025 Conference

Many Voices.
One Community.

Thank you! Any questions?

https://pragseclab.github.io/smudged-fingerprints